Weekly Report

Pingping Shang

2013.8.19~2013.8.25

Work in this week

1. Abstract

We design a system to assist analysts to study the relationship among many labels which are elements of categorical data and how the labels affect the outcome. Many techniques are presented to solve correlation analysis, e.g. scatter plot, parallel coordinates. But we could not find the relationship when it is very complicated and the number of labels is large. Analyst can do much research with our system, e.g. choosing whichever labels they are interested in conveniently, picking up wherever labels and outcome they lead to.

2. Framework

1) Choose labels freely (labeled 1 in figure 1). We can see how many categories every label has and we can decide how to place labels (horizontal or vertical) as to the information about the labels we choose tiled on the canvas (labeled 2 in figure 1) harmoniously.

2) Look for interesting blocks directly (labeled 2 in figure 1). Each block is a group belongs to a categories combination. We use a pixel piece in a block to represent one kind of outcome leaded by the categories combine this block.

3) Watch the relationship among the groups in the MDS view (labeled 4 in figure 1). We select the cluster in the MDS view to do some other research.

4) Set the weight of outcomes(labeled 3 in figure 1). We can set the weight used in the MDS computation of the outcome which we decide to study. The setting view is showed in figure 2.

3. Assess

The system can not provide any proposal about which label might have

significative analysis and the analysts have to try relying on experience.

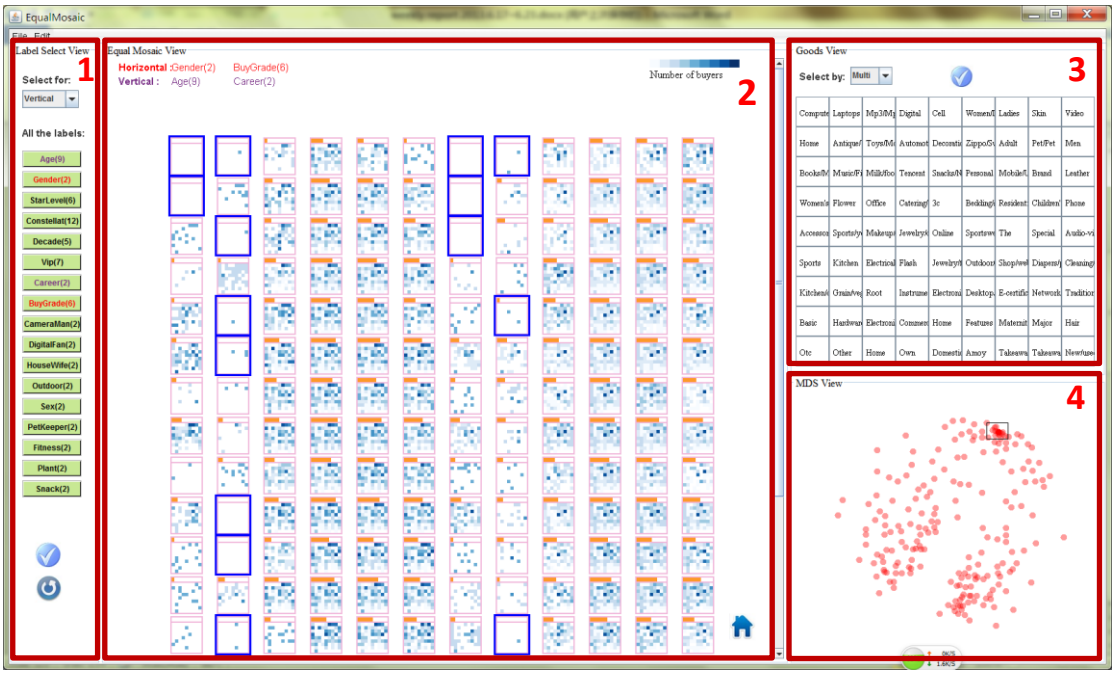

Figure 1

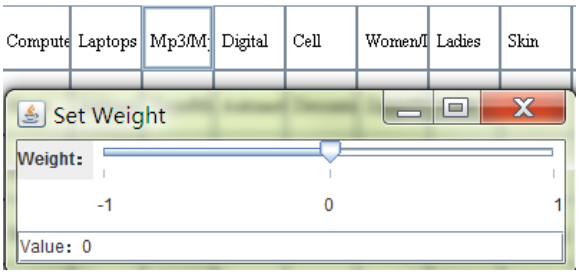

Figure 2